Beyond the identification problem: Modeling age, period, and cohort patterns of mental wellbeing using recent advances in APC analysis

Frans Johannes van Leeuwen, European University Institute

Background & Objective

The life course approach involves three fundamental temporal variables: age, period, and cohort. These variables are not only important for explaining individual outcomes, but also for explaining macro-level differences between age groups, periods, and generations.

The goal of age-period-cohort (APC) analysis is to distinguish between the contributions of each and identify their separate causal effects. But though the distinction between age, period, and cohort effects itself is clear enough, it has proven to involve huge difficulties to distinguish them in practice. The reason is that age, period, and cohort are defined in terms of each other: period (survey year) minus age is cohort (birth year), period minus cohort is age, and age plus cohort is period. This means that none of these three factors can vary independently of the other two, which makes them perfectly collinear. This is not merely a theoretical problem, because if differences are attributed to one of these factors this implies very different explanations than if they were attributed to another.

This issue, which is referred to as the "identification problem", has haunted researchers for more than 50 years. Because of the crucial importance of distinguishing age, period, and cohort effects, many have tried to come up with solutions for this problem. None of these has achieved universal acceptance, and some have even argued that this problem might be impossible to solve at all. However, recent advances have made it possible to isolate trends to an extent previously thought to be impossible. Whereas earlier methods tried to estimate age, period, and cohort trends at the same time using various kinds of complicated mathematics, current approaches such as Fosse and Winship's bounding approach and Luo and Hodges' age-period-cohort interaction (APC-I) model accept that this is impossible and rather focus on the things which can still be done within this limit.

In this study, I apply the most important classical and contemporary methods of APC analysis to the same data about mental wellbeing from one dataset, the Korean Welfare Panel Study (KOWEPS), and investigate to what extent they are able to arrive at the same conclusions with regard to age, period, and cohort patterns in this outcome. In this way, I provide a clear comparison between classical and contemporary approaches to APC analysis which serves as a benchmark to assess to what extent is is actually possible to reach valid conclusions through each method.

Methods

The APC models included in this study are: 1) zero effects / two-factor reduction models, 2) nonlinear transformations, 3) the Intrinsic Estimator (IE), 4) hierarchical age-period-cohort (HAPC) models, 5) the APC-I model, and 6) the bounding approach. A natural way to compare the results of these different models is to place them in the context of Fosse and Winship's bounding approach. These authors proceed from the notion that age, period, and cohort cannot be estimated as separate terms at the same time. However, even though different combinations of the three variables can produce exactly the same outcome estimates, some combinations can be excluded. This is a consequence of the fact that any increase in the size of the linear age and cohort effect can be offset by a similar decrease in the linear period effect in order to achieve the same model estimates and vice versa.

Each model is applied to the same data on mental wellbeing derived from the KOWEPS, a South Korean panel study. As a panel study, it allows looking simultaneously at age, period, and cohort variables, and thus to study how mental health is related to each of them. For the current exposition, the age range was restricted to 18-75, the birth range to later than 1930, while the full period range from 2005 to 2020 was kept. Mental wellbeing is measured in the KOWEPS with the CES-D scale. Scores were positive-standardized based on scores in the first wave in 2005, so that higher scores mean better health and lower scores mean worse mental health, while the average score in this wave was set to 50 and a standard deviation to 10 points. The analysis was conducted only on male respondents, as age, period, and cohort trends may be different for men and women, and including both genders would double the number of analyses, tables, and figures required. Thus, the current analysis includes data on 21,191 individuals, for whom 180,221 distinct observations are available.

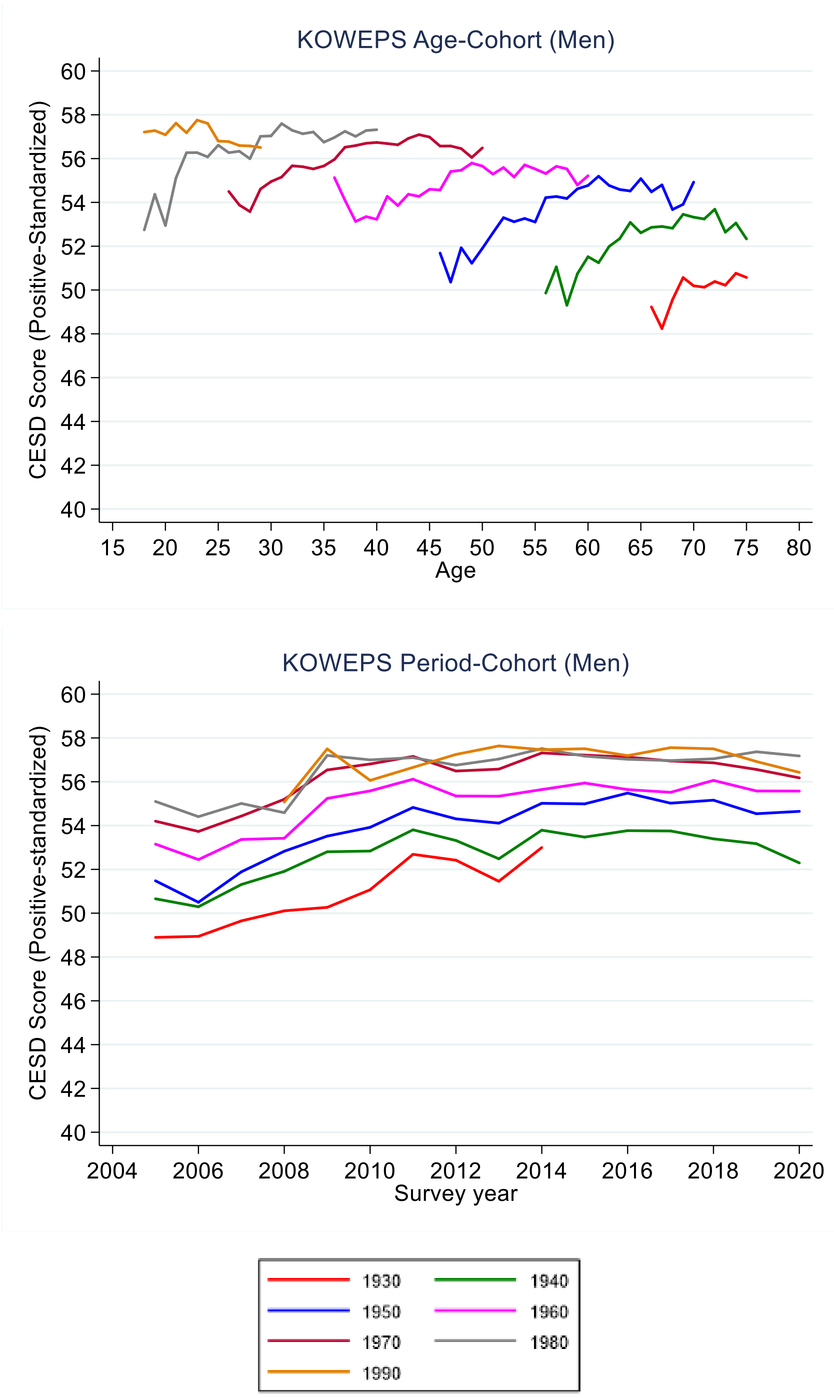

Figure 1. Age-cohort and period-cohort patterns of mental health

Results

Results of the bounding model showed that for the current data there are four possible combinations (Figure 2): 1) a positive age, negative period, and positive cohort effect (upper left), 2) a positive age, positive period, and positive period effect (upper right), 3) a negative age, positive period, and positive cohort effect (middle right), and 4) a negative age, positive period, and negative cohort effect (lower right).

It seems unlikely that there has been a negative period or cohort trend in South Korea for the last 15 years, as economic and social conditions have improved greatly in this country during this time. This means that the possible range of estimates can be further bounded: both period and age effects have to be positive or at least zero. This means that options 1) and 4) can be discarded, leaving us with only options 2) and 3). This accords surprisingly well with the findings from the other models (Table 1), almost all of which postulate either a) a positive age, positive or zero period, and positive cohort trend, or b) a negative age, positive period, and positive or zero cohort trend. Only two models find another combination, namely a negative age, positive period, and negative cohort trend.

Conclusion

Though the identification problem remains a given, these results show that APC methods based on different conceptual and mathematical foundations can actually reach similar substantial findings. Thus, this problem may be less pernicious in practice than it is in theory.

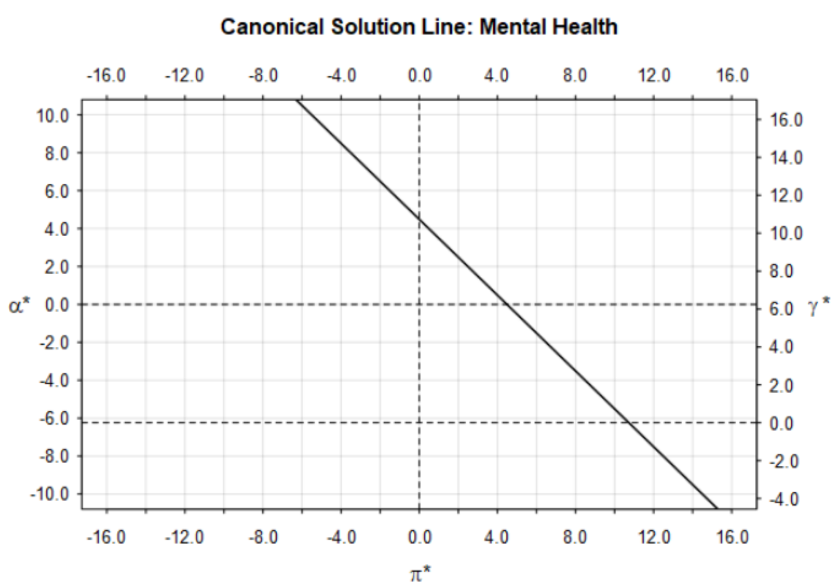

Figure 2. Range of possible APC effect combinations.

Table 1. Summary of model coefficients and correspondence with bounding approach

Model		Variable			Bounded model				BIC
		A	P	C	1	2	3	4	
#1	2F-AC	+	=	+		v			560442
#2	2F-PC	=	+	+		v			560420
#3	1-RE 2F-AC	+	=	+		v			575753
#4	EQ 2005=2020	+	=	+		v			559950
#5	EQ 2005=2006	-	+	-				v	559945
#6	BC (10-year cohort)	-	+	=			v		558961
#7	QT - No LP	+	=	+		v			559893
#8	QT - No LC	-	+	=			v		559893
#9	IE	-	+	=			v		545354
#10	2-RE / HAPC	-	+	=			v		558708
#11	C=A*P	-	+	-				v	560414
#12	P=A*C	+	+	+		v			560388